

ABSTRACT

A system facilitates the deletion of data, such as files, orphaned chunks, and stale replicas. The system may identify a file to be deleted, rename the identified file, permanently delete the renamed file a predetermined amount of time after renaming the identified file as part of a garbage collection process, receive, from the servers, information concerning chunks stored by the servers, and identify, to the servers, ones of the chunks that do not exist possibly due to the permanent deletion of the renamed file. The system may further provide a mapping of file names to chunks, identify chunks, as orphaned chunks, that are not reachable from any of the file names, delete the orphaned chunks, receive, from the servers, information concerning chunks stored by the servers, and identify, to the servers, ones of the chunks that are orphaned chunks. The system may also associate version information with replicas of chunks, identify stale replicas based on the associated version information, delete the stale replicas, receive, from the servers, information concerning replicas stored by the servers, and identify, to the servers, ones of the replicas that are stale replicas.